CLUSTER BASED APPROACH FOR LOAD BALANCING USING K-MEANS IN CLOUD COMPUTING

Utkarsh Urphate [1] | Prof. Santosh Shelke [2]

[1] Student, CSE, SAE Kondhawa, Pune, India -48.

[2] Professor, CSE, SAE Kondhawa, Pune, India -48.

ABSTRACT

Multicloud is a methodology to allocate workload across many computers or other resources over the network links to achieve optimal resource utilization, make the most of throughput, minimum response time, and avoid overload. It presents a load balancing. This research paper proposed an algorithm that focuses on load balancing to reduce the situation of overload or under load on multi-clouds. Task Scheduling algorithms or technique in cloud computing [1]. Efficient task scheduling mechanism should meet user's requirements and improve the resource utilization, so as to enhance the overall performance of the cloud computing environment. It is the prominent means to achieve efficient resource sharing and utilization Cloud Computing is an emerging area in the field of information technology (IT). Load balancing is one of the main challenges in cloud computing. Cloud computing enables a large range of users to access scattered, virtualized hardware and/or software infrastructure over the Internet [2]. The goal of load balancing is to minimize the resource consumption which will further reduce energy consumption and carbon emission rate (with reference to the environment) that is the need of cloud computing. This we will implanting the logs method to the database, using the logs of user activity, to keep track of the data [3]. The future scope is to implementing the cloud on the huge proportion and green computing.

KEYWORDS: Cloud Computing, Load Balancing, Multi Cloud, Load Balancing, Task Scheduling.

INTRODUCTION:

As the scope of cloud scales up, cloud computing service suppliers need handling of gigantic requests. Thus in spite of glorious future of Cloud Computing, many actual problems still essential to be explored for its perfect awareness. One of these concerns is Load balancing. Using multiple components with load balancing instead of a single component may increase reliability and availability through redundancy. Load balancing usually involve dedicated software or hardware, such as a multilayer switch or a Domain

Name System server process. One of the most commonly used applications of load balancing is to provide a single Internet service from multiple servers, Sometimes known as a server farm. Commonly load-balanced systems include popular web sites, large Internet Relay Chat networks, high-bandwidth File Transfer Protocol sites, Network News Transfer Protocol (NNTP) [4] servers Domain Name System (DNS) servers, and databases The latest vision of large distributed computing is Cloud. Cloud based multimedia system (CMS) gained momentum as there is large number of users.

Cloud computing is internet based computing, whereby shared resources software and information are provided to computers. Because when a user uses a personal or professional computer for a great purpose, then they must have some precious files, for which that man is read to invest more invest to protect the file. As the scope of cloud scales up, cloud computing service suppliers need handling of gigantic requests. Thus in spite of glorious future of Cloud Computing, many actual problems still essential to be explored for its perfect.

LITERATURE SURVEY:

1. Tingting Wang; Zhaobin Liu; Yi Chen; Yujie Xu; Xiaoming Dai pro- posed a This paper mentioned regarding the scheduling and load balancing To solve the matter, considering the new characteristics of cloud computing and original adaptive genetic algorithmic program (AGA) a brand new scheduling algorithm supported double effective adaptive algorithm-job spanning time and load balancing genetic algorithm (JLGA) is established [5]. Then compare the performance of JLGA with AGA through simulations.

2. Lock-free multiprocessing solution for LB - X. Liu et al. [4] proposed a lock-free multiprocessing load balancing solution that avoids the use of shared memory in contrast to other multiprocessing load balancing solutions which use shared memory and lock to maintain a user session. It is achieved by modifying Linux kernel. This solution helps in improving the overall performance of load balancer in a multi-core environment by running multiple load-balancing processes in one load balancer.

3. Task Scheduling Algorithm Based on Load Balancing- Y. Fang et al [6] discussed a two-level task scheduling mechanism based on load balancing to meet dynamic requirement of users and obtain high resource utilization. It achieves load balancing by first mapping task to virtual machines and the virtual machines t host resources thereby improving the task response time, resource utilization and overall performance of the cloud computing environment.

4. Scheduling strategy on load balancing of virtual machine resources - J. Hu et al. [7] proposed a scheduling strategy on load balancing of VM resources that uses historical data and current state of the system. This strategy achieves the best load balancing and reduced dynamic migration b using a genetic algorithm. It helps in resolving the issue of load-imbalance and high cost of migration thru achieving better resource utilization

System Architecture / System Overview:

We are building a java, MySQL, data storage (physical cloud), cloud storage (online for the disaster management) based platform project. In which our aim is to provide cluster formation on cloud in which the balanced the load from host to host. A system architecture is a conceptual model that defines the user, application, task, no of clouds ,logs(database logs using for the keep track of the user and user load). The description of architecture is formal description and the representation of a system. It is organized in such a way that supports reasoning about the structures and behaviors of the system. In this system, there are various users. A user can upload a file or a download a file with respect to the his/her requirement. The our task is that using the application layer, we can monitor, scheduling task, maintain the loads on clouds, as per the user is generated, we create a logs in the database, logs with respect to the all activities done by the user, are all written in the text format in the log file. Using the log file we manage the load, managing the less failure rate, increasing the performance, scalability and many more. This project has contain four modules.

1. User & Owner login.

2. Open Application.

3. Upload and download file on servers.

4. Grant Permission.

5. Balance the load using calculating the data based on sever size.

4. Monitoring Logs.

5. Data backup.

Here, In the Multi cloud architecture, our main objective is s to develop an effective load balancing algorithm to maximize or minimize different performance parameters (throughput, latency for example) for the clouds of different sizes (virtual topology depending on the application requirement).Load balancing in cloud computing systems is really a challenge now. Always a distributed solution is required.
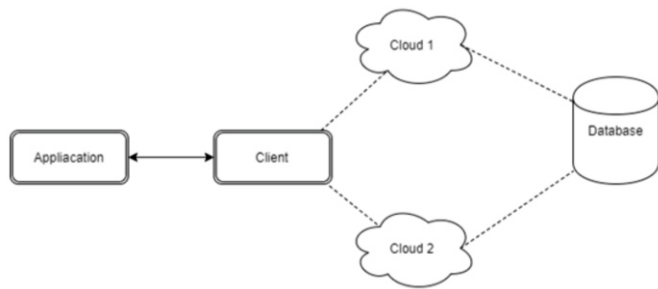
Fig. 1: System Architecture

A load balancing algorithm which is dynamic in nature does not consider the previous state or behavior of the system, that is, it depends on the present behavior of the system. The important things to consider while developing such algorithm are : estimation of load, comparison of load, stability of different system, performance of system, interaction between the nodes, nature of work to be transferred, selecting of nodes and many other ones. This load considered can be in terms of CPU load, amount of memory used, delay or Network load.

SYSTEM ANALYSIS:

System analysis is the act, process or profession of studying an activity typically by mathematical means in order to define purposes or its goals which will discover operations and procedures for accomplishing them efficiently. This section of a system into component pieces for studying how those components work is system analysis. The system analysis is the process of gathering and interpreting facts, diagnosing problems ,and using the information to recommend further improvements to the system. This is the job of system analyst.

The detailed study of various operations performed by a system and their relationships within and outside the system.

RESULT ANALYSIS:

We are assume some factor that should be useful in the result, they are as follows, Confidentiality refers to protect the data from unauthorized users. Cloud computing system offer (data and infrastructure) via public network. Therefore keeping all confidential data of user's secret in the cloud is fundamental requirement which will attract even more users consequently. Data confidentiality can be achieved by using encryption method. Data can be encrypted before placing in cloud provider's server. Only authorized user can access data using decryption method.

Data integrity refer to data cannot be altered by unauthorized person while transferring from or to cloud service provider. Some of the elements used to ensure data integrity such as firewall, communication security management, intrusion detection service, and hash based method.

Table 1: Result

Sr. No.	File Size	Allocate Size	Full Cloud Size	Request Extra Size
1	20MB	20MB	Full	100MB
2	10MB	20MB	Available	Not Need
3	30MB	30MB	Full	50MB

Availability refer to data must be accessible to the authorized user in specified time. The concept of availability guarantee of service, performance and up time. Some of the elements that are used to ensured availability are fault tolerance for data availability such as backup and redundant disk system, reliable and inter Portable security process and network security mechanism

Algorithm Steps:
K-means Algorithm

Input: Data, file size, threshold value

Output: K cluster

Step 1: Determine the size of the file according to the actual situation, according to the size of the data attribute values to assign data to the grid.

Step 2: Calculate each file size quantity and data center

Step 3: Cloud controller receives a new request. Cloud controller queries appropriate node controller/load balancer for next allocation

Step 4: The rest of the cloud center point as input data of obtains K value.

Step 5: After the k value is determined, Choose M file (M>K), their maximum number of file. In the M center points, find K points farthest from each other and those K center points as the initial cluster center of K-means clustering algorithm.

Step 6: According to the initial center point, distributes the data to the nearest cloud.

Step 7: Calculating the average of all the data in each class as the new center.

Step 8: Repeat steps 6 and 7 until the objective function converges.

CONCLUSION:

The approach takes time utilization and resource utilization into consideration and hence results in high signification. The experiments need to be done for processors at various platforms and that reduces the execution time. The main objective of the proposed system is to find out who is trying assigning the task to cloud system and which cloud getting more used. It also ensures that every computing resource is distributed efficiently and fairly. With proper load balancing, resource consumption can be kept to a minimum which will further reduce energy consumption and carbon emission rat which is a dire need of cloud computing. Existing load balancing techniques that have been discussed mainly focus on reducing associated overhead, service response time and improving performance.

REFERENCES:

1. Manel Bourguiba, Member, IEEE, Kamel Haddadou, Member, IEEE, Ines El Korbi, Member, IEEE, and Guy Pujolle, Member, IEEE, "Improving Network I/O Virtualization for Cloud Computing", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 25, NO. 3, MARCH 2014.

2. Veerawali Behal, Anil Kumar,"Cloud Computing: Performance Analysi of Load Balancing Algorithms in Cloud Heterogeneous Environment at 2014 5th International Conference- Confluence The Next Generation Information Technology Summit (Confluence)

3. Sadiku, M.N. Musa, S.M.; Momoh,"Cloud Computing: Opportunitie and Challenges," Potentials, IEEE , vo1.33, no. I , pp.34-36 Jan.-Feb. 2014 doi: 10.1109IMPOT.2013.2279684

4. Kwang Mong Sim Senior Member, IEEE, "Agent-based Cloud Computing" IEEE TRANSACTIONS ON SERVICES COMPUTING,2011.

5. F. M. Aymerich, G. Fenu and S. Sursis, "An Approach to a Clou Computing Network," In Proceedings of the 1st International Conference on the Applications of Digital Information and Web Technologies (ICADIWT), Ostrava, Aug., pp. 113- 118, 2008.

6. A. M. Nakai, E. Madeira, and L. E. Buzato, "Load Balancing for Interne Distributed Services Using Limited Redirection Rates", 5th IEEE Latin-American Symposium on Dependable Computing (LADC), 2011, pages 156-165

7. Xi. Liu, Lei. Pan, Chong-Jun. Wang, and JunYuan. Xie, "A Lock-Free Solution for Load Balancing in Multi-Core Environment", 3rd IEEE International Workshop on Intelligent Systems and Applications (lSA), 2011, pages 1-4.

8. H. Liu, S. Liu, X. Meng, C. Yang, and Y. Zhang, "LBVS: A Loa Balancing Strategy for Virtual Storage", International Conference on Service Sciences (lCSS), IEEE, 2010, pages 257-262

9. Gaochao Xu, Junjie Pang, and Xiaodong Fu, "A Load Balancin Model Based on Cloud Partitioning for the Public Cloud", TSINGHUA SCIENCE AND TECHNOLOGY ISSNl11007-0214l l04/12l lpp34-3 Volume 18, Number 1, February 2013.

10. Y. Fang, F. Wang, and J. Ge, "A Task Scheduling Algorithm Base on Load Balancing in Cloud Computing", Web Information Systems and Mining, Lecture Notes in Computer Science, Vol. 6318, 2010, pages 271-277

11. Klaithem Al Nuaimi, Nader Mohamed, Mariam Al Nuaimi and Jameel Al-Jaroodi, "A Survey of Load Balancing in Cloud Computing: Challenge and Algorithms", 2012 IEEE Second Symposium on Networ Cloud Computing and Applications.

12. Vijay Varadharajan, Senior Member, IEEE, an Udaya Tupakula, Member, IEEE Security as a Service Model for Clou Environment March 2014.

13. Jun-Feng Zhao Jian-Tao Zhou Strategies and Methods for Cloud Migration April 2014.

14. Qiao Yan, F. Richard Yu, Senior Member, IEEE, Qingxiang Gong, and Jian qiang Li Software-Dened Networking (SDN) and Distributed Denial of Ser- vice (DDoS) At tacks in Cloud Computing Environments: A Survey, Some Research Issues and Challenges 2016.

15. Qiao Yan, F. Richard Yu, Senior Member, IEEE, Qingxiang Gong, and Jian-qiang Li Software-Dened Networking (SDN) and Distributed Denial of Ser- vice (DDoS) Attacks in Cloud Computing Environments: A Survey, Some Research Issues and Challenges 2016.

16. Hussein Khalid Abd-alrazzq1, Mohammad S. Ibrahim2 and Omar Abdul Rahman Dawood 3 \Secure Internet Voting System based on Public Key Kerberos", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 3, March 2012.